



Refractive index tomography with a physics-based optical neural network

DELONG YANG,¹ SHAOHUI ZHANG,^{1,2,3} CHUANJIAN ZHENG,¹
GUOCHENG ZHOU,¹ YAO HU,¹  AND QUN HAO^{1,2,4}

¹*School of Optics and Photonics, Beijing Institute of Technology, Beijing 100081, China*

²*Yangtze Delta Region Academy of Beijing Institute of Technology, China*

³*zhangshaohui@bit.edu.cn*

⁴*qhao@bit.edu.cn*

Abstract: The non-interference three-dimensional refractive index (RI) tomography has attracted extensive attention in the life science field for its simple system implementation and robust imaging performance. However, the complexity inherent in the physical propagation process poses significant challenges when the sample under study deviates from the weak scattering approximation. Such conditions complicate the task of achieving global optimization with conventional algorithms, rendering the reconstruction process both time-consuming and potentially ineffective. To address such limitations, this paper proposes an untrained multi-slice neural network (MSNN) with an optical structure, in which each layer has a clear corresponding physical meaning according to the beam propagation model. The network does not require pre-training and performs good generalization and can be recovered through the optimization of a set of intensity images. Concurrently, MSNN can calibrate the intensity of different illumination by learnable parameters, and the multiple backscattering effects have also been taken into consideration by integrating a "scattering attenuation layer" between adjacent "RI" layers in the MSNN. Both simulations and experiments have been conducted carefully to demonstrate the effectiveness and feasibility of the proposed method. Experimental results reveal that MSNN can enhance clarity with increased efficiency in RI tomography. The implementation of MSNN introduces a novel paradigm for RI tomography.

© 2023 Optica Publishing Group under the terms of the [Optica Open Access Publishing Agreement](#)

1. Introduction

Remarkable three-dimensional (3D) visualization of biological processes is of paramount importance for biological research. This is particularly challenging due to the transparent or semi-transparent characteristics of numerous biological cells. Direct microscopic observation often results in a low imaging contrast, missing much information. For improving the imaging contrast and signal-to-noise ratio (SNR) of 3D distributed bio-samples, fluorescence microscopy is one of the most widespread solutions [1–4], and has inspired several representative super-resolution methods, such as Stimulated emission depletion microscopy (STED) [5,6], Stochastic optical reconstruction microscopy (STORM) [7,8], Photoactivation localization microscopy (PALM) [9], et al. Combined with necessary mechanical position scanning techniques, confocal designs [10], multi-photon microscopy [11] and light sheet microscopy [12] can be implemented, fluorescence microscopy enables the achievement of high-resolution 3D imaging of typical biological samples, with both high signal-to-noise ratio (SNR) and spatial resolution. Nevertheless, fluorescence microscopy necessitates exogenous labeling. And since some cellular characteristics or structures may not be detectable using fluorescence, it is unable to probe the total internal structure of biological samples.

Unlike fluorescent labeling methods, phase imaging by measuring the change in the optical path difference of light is another approach that can effectively observe the structure information of transparent and semi-transparent samples. Combining with angularly-resolved measurements

and tomography reconstruction algorithms, 3D refractive index (RI) imaging can be achieved, termed optical projection tomography [13–15]. Although optical tomography can produce high-quality 3D images, the interferometry-based method for obtaining phase distribution under each angular projection is highly sensitive to environmental conditions and prone to speckle noise. Therefore the corresponding system construction cost is high, and the system is less robust.

Besides interferometry, phase retrieval with intensity images utilizing optimization algorithms is also a widely used phase imaging technique. Such methods are not affected by laser speckle noise and are robust to environmental disturbances, so the implementation and operation costs of the systems are also relatively low. Nevertheless, phase imaging from intensity images is not trouble-free. The ability of an optimization algorithm to find a satisfactory global solution depends on the inverse problem's complexity, the physical constraints used, and the chosen initial value. RI tomography from intensity images, referred to as intensity diffractive tomography (IDT) [16–23], is a complex inverse problem involving a large number of unknown parameters and complex multiple scattering physical processes. Conventional IDT approaches, whether they utilize the multi-slice model in the space domain or Ewald's diffractive sphere model in the Fourier domain, have the problems of being seriously time-consuming, slow convergence, and easy to fall into local minima traps [24].

In recent years, many researchers have been actively exploring appropriate deep-learning methods to achieve better performance in computational imaging including tomography. While deep learning has been demonstrated to improve the speed and quality of tomography across various modalities [25–31], conventional deep learning approaches remain unsuitable for certain measurement applications, especially RI tomography of biological samples. Since the conventional deep learning methods are mostly data-driven schemes, the inferential capabilities of the convolution neural networks (CNNs) are from the "experience" gleaned from huge datasets [32–35]. However, constructing a comprehensive dataset for biological tomography remains a significant challenge. Although a physics-based simulator has been proposed to generate training datasets for CNNs in biological tomography, the results are still limited by the mismatch between simulation and experimental data. [36]. Moreover, the accuracy of results from conventional CNNs, when applied to completely novel samples, remains questionable. [37,38].

To address the aforementioned challenges, researchers have proposed the use of neural networks, integrated with physical models, to enhance the accuracy of imaging. In 2020, Horstmeyer Roarke proposed the deep prior diffraction tomography [39] which employs neural networks as the phase retrieval algorithm and uses a light scattering model as the physical verification to improve the authenticity of the imaging. Guohai Situ practiced a similar scheme in coherent diffraction imaging [40] in the same year. In 2022, Ulugbek S. Kamilov proposed the deep continuous artefact-free RI field, which uses the neural field network to implement phase retrieval. The sparse representation capability of neural field networks significantly reduces memory usage and addresses the issue of missing cones in intensity diffraction tomography [41]. However, these physics-based methods still employ conventional neural network structures as inverse operation networks. The optimization process essentially involves searching for the optimal inverse problem model within the vast parameter space of CNNs, a process that remains time-consuming.

In this work, we remodel the structure of conventional neural networks [42] to incorporate both the beam propagation model and the scattering model, as opposed to solely relying on conventional architectures. The beam propagation method [43–46], dividing the RI of the sample into layers, is adapted within the structure of our neural network [47–49]. The beam propagation model imposes constraints on the backward gradient and helps verify the authenticity of the imaging results. The neural network we propose, which incorporates an optical structure, is designated as the multi-slice neural network(MSNN).

To enhance imaging quality, we incorporate dark field raw data into the optimization [50]. We extend the MSNN connection for learnable parameters [51] to calibrate the intensity between

different LEDs, reducing the requirements for acquisition hardware and the need for image preprocessing.

The exceptional fitting capability of MSNN even offers the potential to take the higher-order scattering within the sample into phase retrieval. We take into account the backscattering field as predicted by the scattering model. The multiple backscattering effects have also been taken into consideration by integrating a "scattering attenuation layer" between adjacent "RI" layers in the MSNN. This approach ensures that our network architecture more accurately mirrors the actual process of light field propagation within the sample. We demonstrate that the implementation of adaptive recovery of backscatter intensity using MSNN during *C. elegans* imaging can enhance the quality of the images.

2. Method

2.1. Principle of the multi-slice neural network (MSNN)

Multi-slice beam propagation method separates 3D objects into a series of thin layers, where the light wave through the sample is modeled by sequential layer-to-layer propagation of the light field. The structure of the beam propagation model is very similar to the hierarchical structure of the Deep Neural Networks (DNNs). And there are many optimizers [52] for DNNs to quickly converge to the global optimum. In MSNN, We model the m th layer of the 3D object as $L_m(r) = n(r, m\Delta z) - n_{media}$. Mathematically, the diffraction propagation in MSNN can be recursively written as:

$$P(r, \Delta z) = \exp(j2\pi\Delta z(\frac{n_{media}^2}{\lambda} - \|u\|^2)^{1/2}) \quad (1)$$

$$t_m(r, \Delta z) = \exp(j2\pi\Delta z \frac{n(r, m\Delta z) - n_{media}}{\lambda}) \quad (2)$$

$$U_{m+1}(r) = t_m(r, \Delta z) \cdot \mathcal{F}^{-1}\{P(r, \Delta z) \cdot \mathcal{F}\{U_m(r)\}\} \quad (3)$$

where $P(r, \Delta z)$ denotes the angular spectrum diffraction equation that propagates a light field by distance Δz , r denotes the 2D spatial position vector, u denotes the 2D spatial frequency space coordinates vector, $t_m(r, \Delta z)$ denotes the phase modulation by the m th layer, U_m and U_{m+1} are the input and output light field of m th layer in MSNN. The boundary condition to initialize the recursively Eq. (3) is the incident plane wave illuminating the sample, $U_0(r) = \exp(jk_{illu}r)$ where k_{illu} is the illumination wave vector at a particular angle. The experiment system and the principle of MSNN are shown in Fig. 1.

Simultaneously, based on the 1st Born approximation, the scattering model that suggests the light field through the 3D sample including the incident field and scattered field stimulated by the 3D scattering potential $V(r, z) = k_0^2(n_{media}^2 - n^2(r, z))$. The total light field through the 3D sample can be expressed as:

$$U_{total}(r, z) = U_{in}(r, z) + U_s^{(born)}(r, z) \quad (4)$$

$$U_{in}(r, z) = \mathcal{F}^{-1}\{P(r, \Delta z) \cdot \mathcal{F}\{U_{total}(r, z - \Delta z)\}\} \quad (5)$$

$$U_s^{(born)}(r, z) = \iiint G(r - r', \Delta z) U_{total}(r', z - \Delta z) V(r', z) d^3 r' \quad (6)$$

$U_s^{(born)}(r, z)$ denotes the scattered field under the 1st Born approximation. The beam propagation model and the 1st Born approximation scattering model both consider the propagation of the input field according to the angular spectrum diffraction method, but they are different in the prediction of the scattering field. And the previous work has verified that both models are effective for 3D imaging of weakly scattered thick samples [44,53]. In scattering model, Eq. (6) suggests that the superposition field of spherical waves computed by the Green function is actually the total

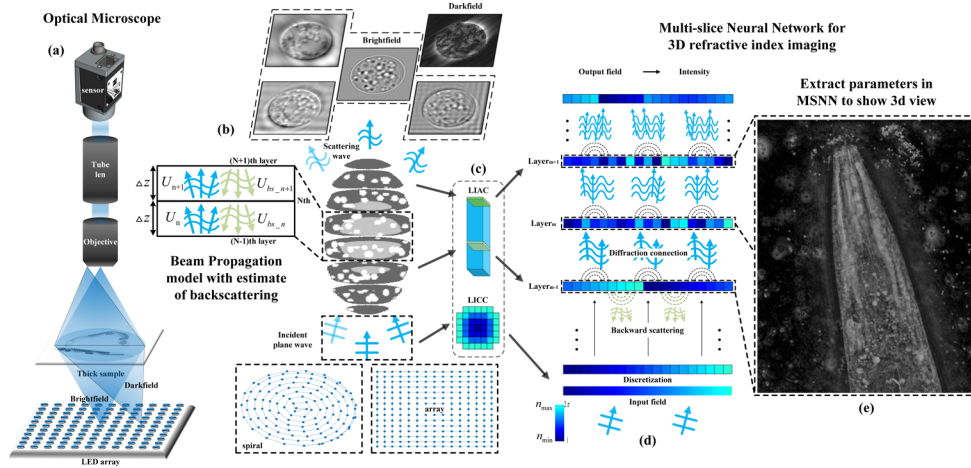


Fig. 1. The principle of MSNN. (a) The optical system of the RI tomography. (b) The beam propagation model with estimate of backscattering. The lighting mode can be selected array lighting or spiral lighting. The bright-field and dark-field images are simulation captured images of a cell phantom. (c) The backscattering of different layers in the sample is predicted in MSNN with Learnable Intensity Attenuation Coefficient (LIAC). The inconsistency in intensity of various incident plane waves is corrected by Learnable Intensity Compensation Coefficient (LICC) in MSNN. (d) The structure of MSNN. After the optimization, we will extract the internal parameters of the neural network as the RI reconstruction results of the sample. (e) The reconstruction result of the head of a *C.elegan*. The image is a two-dimensional projection of the reconstruction result on the x-y plane.

scattering field including the forward scattering field and backscattering field, and we proved that when the layer is thin enough, the forward field and backscattering field can be approximate have same distribution in phase but different in intensity. We transferred the conclusion to the beam propagation model and got better imaging results in the experiments. In the MSNN, we divide the 3D RI of samples into slices, assuming that each slice is thin enough to disregard the scattering within it. The scattering model suggests that the phases of the forward and backscattering fields are identical [53] when the slices are extremely thin:

$$U_s^{(born)}(r, z) = \tilde{G}(u, \Delta z) \iint \text{sinc}(\nabla(u') + \nabla(u)) \Delta z \tilde{U}_n(u') \tilde{V}_n(u - u') \Delta z d^2 u' \quad (7)$$

$$U_{bs}^{(born)}(r, z) = \tilde{G}(u, \Delta z) \iint \text{sinc}(\nabla(u') - \nabla(u)) \Delta z \tilde{U}_n(u') \tilde{V}_n(u - u') \Delta z d^2 u' \quad (8)$$

$$\nabla(u) = \left(\sqrt{\frac{n_m}{\lambda}} - \|u\|^2 \right) \quad (9)$$

where Δz is the thickness of the slice, allowing us to approximate the *sinc* function to one. By applying this conclusion to the beam propagation model, we can consider that the backscattering field attenuates the forward scattering field, although the intensity of this attenuation remains unknown. Based on the simulation experiment in Chen's paper [53] (using a simulation cell model with a scale and RI bandwidth similar to the *C.elegans* sample used in this paper, and the cell model is more uniform and sparser than the *C.elegans*), A rough estimate would lead to an intensity prediction difference of 0%-16% ($\frac{0}{1.8} - \frac{0.08}{0.5}$) in different areas of the image. We believe that this may lead to model degradation in the later stages of optimization. So we employ the beam propagation model to predict the backscattering field and attenuate the incident field

in the layer of the forward propagation in MSNN, we set learnable parameters ω_m to predict the intensity of backscattering field to attenuate forward field, the parameters ω_m which we called Learnable Intensity Attenuation Coefficient(LIAC) are optimized by MSNN. The forward propagation in MSNN can be expressed as:

$$U_{m+1}(r) = t_m(r, \Delta z) \cdot \mathcal{F}^{-1}\{P(r, \Delta z) \cdot \mathcal{F}\{U_m\}\} / \omega_m \quad (10)$$

The exit electric-field, $U_M(r)$, accounts for the light field passing through the sample. When the imaging system is focus at the center of the sample, we need to refocus the exit light field $U_M(r)$ to the center of the sample, where M denotes the total layers of the model. The final light field and intensity distributions at the image plane are:

$$U_{predict}(r) = \mathcal{F}^{-1}\{C(u, k_0) \cdot P(r, -\frac{M\Delta z}{2}) \cdot \mathcal{F}\{U_M\}\} \quad (11)$$

$$I_{predict}(r) = |U_{predict}(r)|^2 \quad (12)$$

where $C(u, k_0)$ denotes the coherent transfer function of the system. The light field captured by the objective $U_{predict}(r)$ accounts for the accumulation of the diffraction and multiple-scattering processes(including the backscattering) that occurred during optical propagation through the $(-\frac{M\Delta z}{2}, \frac{M\Delta z}{2})$ around the focal plane.

2.2. Adaptive intensity calibration between different input fields in MSNN

In practical experiments, to enhance lateral and axial resolution and improve the success rate of 3D imaging, we need to provide plane waves k_i from various angles. When $k_i > NA_{obj}$, the camera captures dark-field images which require longer exposure times to enhance the signal-to-noise ratio. During the image capture process, we set distinct exposure times for different illumination angles. However, the relationship between exposure time length and sensor responsiveness is not strictly linear. Despite raw data preprocessing according to exposure times, this discrepancy could potentially introduce complications in solving the inverse problem.

Furthermore, due to limitations in the accuracy of the experimental system, the intensity of the high-order diffracted input light field (provides illumination for darkfield) - which is used as an initial condition of MSNN - may be inaccurate. To address discrepancies in the response gray value in raw data caused by varying sensor exposure time settings at different illumination angles, and to compensate for the intensity of higher-order diffracted light, we introduce learnable parameters γ_n for each illumination angle. Consequently, the incident wave can be expressed as follows:

$$U_0^n(r) = \gamma_n \cdot \exp(jk_{illu}r) \quad (13)$$

We define γ_n as the Learnable Intensity Compensation Coefficient (LICC) for the n th incident plane wave. In our model, LEDs in different positions each have independent coefficients. These coefficients can be learned and optimized within the Multi-Slice Neural Network(MSNN) during imaging. Consequently, the MSNN identifies the optimal set of LICCs ($\gamma_1 \cdots, \gamma_n \cdots, \gamma_N$) for the LEDs, thereby enhancing the 3D imaging process.

2.3. RI tomography with MSNN

MSNN is employed to solve the following optimization with an objective consisting of a measurement loss \mathcal{L} and regularizer \mathcal{R} :

$$\operatorname{argmin}\{\mathcal{L}(\operatorname{MSNN}(k_{illu}), I_{gt}(r)) + \beta \cdot \mathcal{R}(\operatorname{MSNN})\} \quad (14)$$

where \mathcal{L} denotes the L1 loss. \mathcal{R} denotes the regularization term. We use the 3D total variation (TV) norm as the regularization loss within the MSNN. The parameter β is manually adjusted

to optimize the strength of the regularization, with a larger punishment in the axial direction typically yielding better results. For the MSNN updates, we employ the Adam optimizer [54], which adaptively adjusts the learning rate based on raw data. This process for 3D Intensity-based RI imaging with MSNN is summarized in Algorithm 1.

Algorithm 1. 3D intensity-based RI imaging with MSNN

```

1 Input:The illuminate  $k_{illu}$  of  $N$  LEDs,  $M$  layers of imaging axial size  $M\Delta z$ .
2 Data:Measured intensities  $\{I_{measure}^n(r)\}_{n=1}^N$ .
3 Hyperparameters:The parameters of the microscope system, the step size of
  optimization  $\alpha$ , the TV regularization coefficient  $\beta$ , the learning rate of LIAC  $\omega_m$  and
  LICC  $\gamma_n$ , max number of iteration  $I$ 
4 Initialization:The RI in the layer of MSNN  $\{L_m(r)\}_{m=1}^M = 0$ .
5 Return:3D RI of the sample.
6 for  $i = 1 : I$  do
7   for  $n = 1 : N$  do
8      $k_{illu} = k_n, I_{gt}(r) = I_{measure}^n(r)$ 
9      $U_0(r) = \gamma_n \cdot \exp(jk_{illu}r)$ 
10    for  $m = 1 : M$  do
11       $L_m(r) \leftarrow$  mth layer of MSNN
12       $t_m(r, \Delta z) \leftarrow \exp(j2\pi\Delta z \frac{L_m(r)}{\lambda})$ 
13       $U_{m+1}(r) \leftarrow t_m(r, \Delta z) \cdot \mathcal{F}^{-1}\{P(r, \Delta z) \cdot \mathcal{F}\{U_m(r)\}\} / \omega_m$ 
14    end
15     $I_{predict}(r) \leftarrow |U_M(r)\mathcal{F}^{-1}\{C(u, k_0) \cdot P(r, -\frac{M\Delta z}{2}) \cdot \mathcal{F}\{U_M\}\}|$ 
16     $Loss \leftarrow L_1(I_{predict}(r), I_{gt}(r))$ 
17     $Loss.autograd().backward()$  by optimizer  $Adam(\alpha)$ 
18  end
19   $RI \leftarrow$  layers of MSNN
20   $RI_{reg} \leftarrow TV_{3D}(RI, \beta)$ 
21   $MSNN \leftarrow RI_{reg}$ 
22 end

```

In addition, there are still some tricks that help us optimally configure the network structure to accelerate tomographic reconstruction. For example, users can determine the number of network layers, M , based on the system's minimum axial resolution Δz and imaging depth Z . Alternatively, the number of network layers, M , can be determined based on the minimum axial resolution Δz corresponding to the lateral scale range of the sample. As mentioned in section 4.3 that follows. As for the selection of hyperparameters for MSNN, The optimal hyperparameter (Generally, the adjustment is the learning rate for LICC) can be obtained by monitoring the first few iterations or performing a small number of iterative reconstructions in a localized area.

During the optimization process, the number of iterations is influenced by the sample and the settings of MSNN, including numbers of layer, imaging depth, and the internal multiple scattering of sample. Generally speaking, users can opt to halt once the loss shows no significant reduction, achieving satisfactory imaging quality in such cases.

3. Simulations

To assess its capability for quantitative 3D RI reconstruction, we initially create a pure phase phantom composed of microspheres of varying radius. These microspheres have a RI of 1.59,

while the medium has an RI of 1.56. The ideal microspheres, with diameters ranging from 1 to 4 μm and an RI of 1.59, are immersed in a medium matching the RI of 1.56. The numerical aperture (NA) of both the objective (Magnification 40X) and the illumination in the simulation is 0.7. We sequentially illuminate angle-varied plane waves, with a center wavelength of 532 nm, through the sample and generate 225 captured intensity images. Each image contains 100×100 pixels, with a sensor pixel size of $4\mu\text{m}$. Figure 2 shows the 3D imaging results of the microspheres, with MSNN producing a satisfactory quantitative 3D RI. However, larger radius microspheres demonstrate larger axial artifacts, attributable to the problem of low-frequency missing cones in the frequency domain.

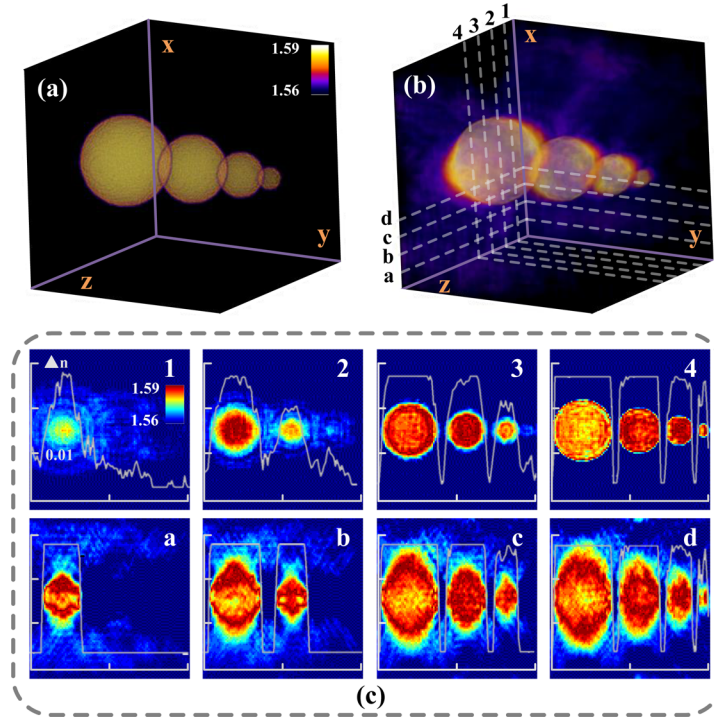


Fig. 2. The reconstruction results for 4 microspheres of varying sizes (a) This panel depicts the 3D ground truth of the microspheres, with radius of $4\mu\text{m}$, $3\mu\text{m}$, $2\mu\text{m}$ and $1\mu\text{m}$. (b) The 3D reconstruction results of the microspheres. (c) The reconstructions at depths of $1\mu\text{m}$, $2\mu\text{m}$, $3\mu\text{m}$ and $4\mu\text{m}$ in both the x-y and y-z planes.

Subsequently, we create a synthetic cell phantom containing intricate details for 3D RI recovery, with an RI that ranges from 1.33 to 1.38. Figure 3 depicts the imaging results of the cell phantom by MSNN, using 225 measurements. Even though it is subjected to the missing cone problem and compromises axial resolution for low spatial frequencies, MSNN manages to reconstruct a satisfactory quantitative 3D RI. Simulation was performed with PyTorch running on a desktop computer equipped with Intel Core i5-10500 CPU at 4GHz 32GB RAM CPU and NVIDIA's GeForce GTX 3090 GPU. The total imaging time to complete 10 iterations for a volume of $100 \times 100 \times 100$ voxels was 40 seconds.

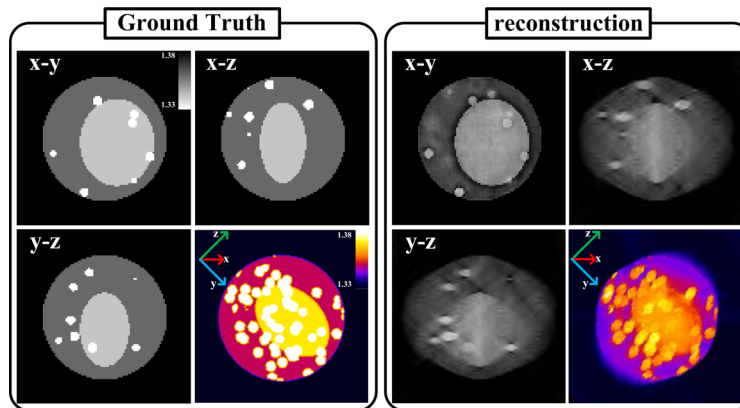


Fig. 3. The middle slice of x-y,x-z,y-z plane of the ground truth and imaging with MSNN. And the 3D projection views in a tilt angle are present.

4. Quantitative verification

4.1. System setup

To demonstrate the improvement of MSNN, we use a two-slice test sample consisting of two resolution targets, one placed above the focal plane and the other axially spaced apart and rotated relative to the first one. We build a microscope system as shown in Fig. 1 with a 4x objective ($NA \approx 0.1$). We use a LED array (15×15) for varied-angle illuminations, and the distance between adjacent LEDs is 5mm. The LED array is located 100mm from the sample plane and emits a light with a wavelength of 473nm and a bandwidth of 20nm. The LED array is controlled by an stm32 microcontroller and is synchronized with the camera to scan through the LEDs at camera-limited speeds. Our camera is capable of 50 frames per second at full frame (2448×2048 pixels, pixel size $3.45 \mu\text{m}$) and with 8-bit data. However, for dark-field images, we opt for longer exposure times. The raw data, constrained to a small 360×360 pixel resolution region, is used to recover a high-resolution complex field with a 4X increase in pixel quantity. The computing platform is aligned with the simulation. The total imaging time to complete 50 iterations of a $720 \times 720 \times 2$ voxels volume was 1 minutes.

4.2. Experiment results

The low-resolution raw images captured consist of 9 brightfield and 216 darkfield images. Since the USAFchart serves as an intensity modulation sample, we employ a complex RI matrix to represent the object. The imaginary parts of the complex RI sections, representing the absorption of the light field at different depths within the sample, are shown in Fig. 4(d). MSNN successfully isolates the intensity modulation information at various depths from the raw data.

However, without the LICC, MSNN fails to achieve the theoretical lateral resolution even after 500 optimization iterations, and the imaging quality across depth sections is unsatisfactory. The compromised quality also suggests the presence of system deviations, such as pupil aberration, illumination deviation, and exposure time differences [55–57].

When using the LICC, we manage to reconstruct the sample at an enhanced resolution in as few as 50 iterations. With extended optimization, the lateral resolution improves steadily. However, our experiments show that the optimal learning rate for LICC varies across training iterations. As shown in Fig. 4(c), sudden drops in loss occur at marked positions within different loss curves, resulting in a degradation of imaging. Therefore, it's necessary to halt optimization

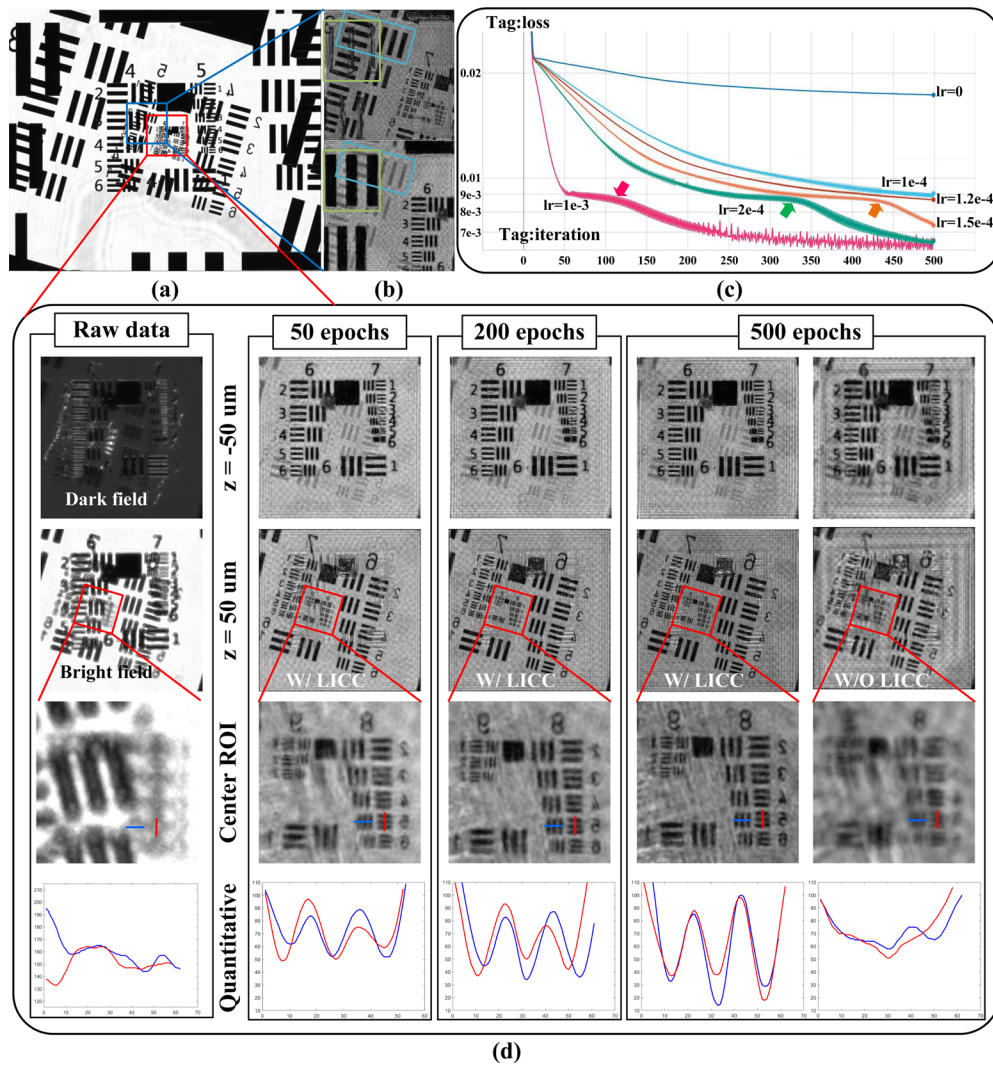


Fig. 4. Absorption imaging of a two-slice sample consisting of two resolution targets placed at different depths (-50 and $50 \mu\text{m}$), with one rotated laterally with respect to the other. (a) Low-resolution full FoV image from $4\times 0.1\text{NA}$ objective. (b) The imaging for 5-3 ($12.41\mu\text{m}$) and 4-3 ($24.80\mu\text{m}$) in two slice. (c) The loss curve for different learning rates of light calibration. The overfittings during the optimization are labeled, and the training should be stopped before the overfitting. (d) The captured raw data and the reconstructed results with different learning rates of LICC

early by observing the loss trend. The early stop point loss within the different curves further supports the notion that LICC improves imaging quality.

4.3. Analysis of resolution

4.3.1. Formulation of Ewald sphere

The three-dimensional Fourier diffraction theorem suggests the lateral resolution and axial resolution in the intensity diffractive tomography (IDT) system are not independent, but the structures in the biomedical sample with large lateral size usually have a large axial size, we

usually ignored studying the relationship between the axial resolution and lateral resolution in our experiment system.

To analyze both lateral resolution and axial resolution theoretically, we use the three dimensional (3D) coherent transfer function (CTF) of the imaging system. As illustrated in Fig. 5(a), the 3D spatial frequency(k -space) is formed by the 3D Fourier transform of the object's RI [58,59]. This spherical k -space representation, known as the Ewald sphere, has a radius determined by the wave vector $k_0 = \frac{2\pi}{\lambda}$.

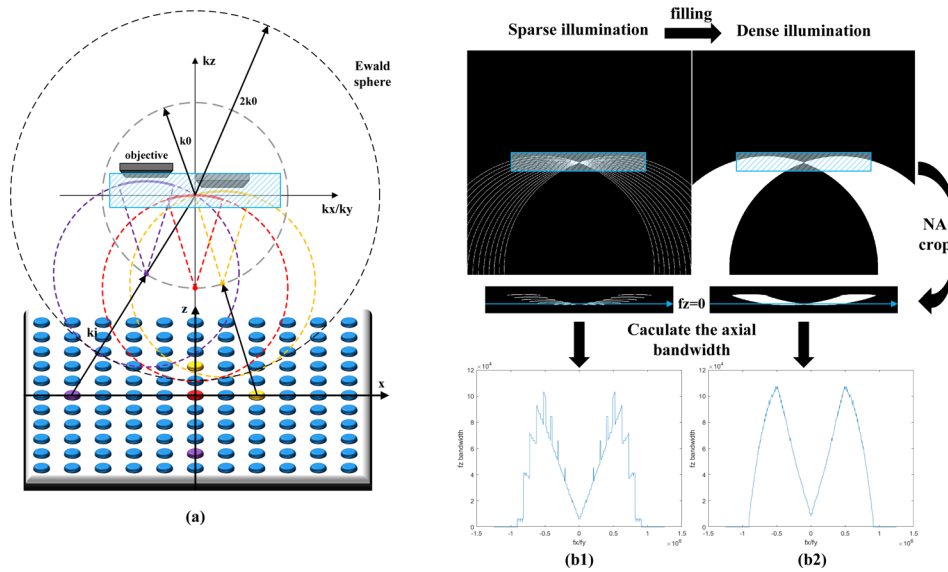


Fig. 5. Fourier diffraction theorem in finite-aperture optical systems. (a)The 3D sample is illuminated by plane waves from various angles. The forward scattering wave captured by the aperture of the objective with fixed NA. (b1)The bandpass on the Ewald sphere and the axial bandwidth in different lateral resolutions by the sparse illumination of (a). (b2)The bandpass on the Ewald sphere and the axial bandwidth in different lateral resolutions by the dense illumination

By sequentially activating each light-emitting-diode(LED) at different positions on the LED array, which in turn provides plane waves k_i at varying angles, we can explore different regions of the k -space. As k_i changes with the illumination angle, the maximum probed k -space falls within a spherical shell with radius k_0 . The center of this shell shifts along a second spherical shell(of the same radius k_0) determined by the incident angle of the plane wave k_i (as depicted by the gray circle in Fig. 5(a)).

However, in actual experiment systems, The Fourier diffraction theorem suggests that, limited by illumination and microscopy, only partial spherical cap bounded by the generalized aperture can be probed. Illuminating the object at different angles will shift different regions of the object's k -space into a fixed microscope objective lens with fixed numerical aperture (NA). Ultimately, only a portion of the Ewald sphere can be reconstructed. [17,20].

We employ matrices to digitally simulate the 3D CTF based on our system setup. The 3D k -space bandwidth in the x - z section(which is equivalent to the bandwidth in the y - z section) is depicted in Fig. 5(b). From this, we can compute the theoretical axial resolution by assessing the k -space bandwidth of the x - z region. The results suggest that different lateral resolutions correspond to varying axial resolutions. This implies that the imaging quality for line pairs on a USAF chart will differ across the depth(z) section due to these resolution disparities.

4.3.2. Analysis of axial resolution for the system of quantitative experiment

The LED array provides various illumination angles to improve the lateral resolution [60,61]. As expected, the lateral bandwidth Δfx is determined by the sum of objective $NA(NA_{obj})$ and illuminate $NA(NA_{illu})$:

$$\Delta fx = \frac{2(NA_{obj} + NA_{illu})}{\lambda} \quad (15)$$

According to Eq. (15), The theoretical lateral resolution is $0.821 \mu\text{m}$, indicating that the clearest line pair on the USAF chart would be between 9-2 and 9-3. The axial resolution, however, does not follow this trend. Its bandwidth, Δfz , is neither that of the objective nor is it that would result from using an objective having the sum of the two NA. Instead, it is somewhere in between the NA of the objective and the sum of two NAs (NA_{obj} and NA_{illu}), and can be estimated with a formula:

$$\Delta fz = \frac{2 - \sqrt{1 - NA_{obj}^2} - \sqrt{1 - NA_{illu}^2}}{\lambda} \quad (16)$$

Nonetheless, the calculated result using this formula represents the maximum axial resolution achievable across various lateral resolutions. The 3D Fourier diffraction theorem [62] suggests the lateral resolution and axial resolution in the Fourier Ptychography system for RI tomography are not independent [63]. We propose a digital simulation method for more accurate analysis of axial resolution in microscopy systems. In the USAFchart illuminated by our system, the 5-3 line has $44 \mu\text{m}$ theoretical maximum axial resolution, and the 4-3 line pair has $70 \mu\text{m}$ theoretical maximum axial resolution. In the green ROI of Fig. 4, the information regarding the separation of the 4-3 line pairs at an axial distance of $100 \mu\text{m}$ is distinct in different depth sections. In order to further analyze the difference between the axial resolution of our system and the theoretical limit, we narrowed the distance of USAFchart for further experiments.

Since the lines with the different lateral sizes in USAFchart have the same axial size, the stacked resolution targets provide a convenient way to experimentally characterize lateral resolution at multiple depths. At the same time, we can transform the various linewidth which represents various lateral resolutions into the spectrum and view the reconstruction of depth(z) section to analyze the axial resolutions.

According to the axial bandwidth in Fig. 5(b1) which uses sparse illumination, the maximum axial resolution is $9.6 \mu\text{m}$ in $1.9 \mu\text{m}$ lateral resolution (8-1 line pair). The 5-3 line pair with $12.41 \mu\text{m}$ in Fig. 5(b) has $57 \mu\text{m}$ axial resolution, the 4-3 line pair with $24.8 \mu\text{m}$ in Fig. 5(b) has $97 \mu\text{m}$ axial resolution. In dense illumination, the maximum axial resolution is $9.3 \mu\text{m}$ in $2 \mu\text{m}$ lateral resolution. The 5-3 line has $44 \mu\text{m}$ axial resolution, and the 4-3 line pair has $70 \mu\text{m}$ axial resolution. The difference between axial resolutions in low-resolution lines is mainly caused by numerical simulation errors (The numerical error can be reduced by expanding the matrix scale), but the analysis still can be used as a reference. The actual axial resolution should lie between the above two sets of data, taking into account the density of the illumination.

To confirm whether the axial resolution of a 4-3 or 5-3 pair aligns with the theoretical analysis value, we present the reconstruction results of a pair of USAF charts, reconstructed by MSNN, with a placement spacing of $54 \mu\text{m}$. As observed in Fig. 6, ROI① with smaller lateral sizes demonstrate better layering compared to ROI② and ROI③, which have larger lateral sizes. And due to the axial bandwidth not being symmetrical, ROI② and ROI③, which have the same lateral size, also exhibit different layering effects in planes at varying depths.

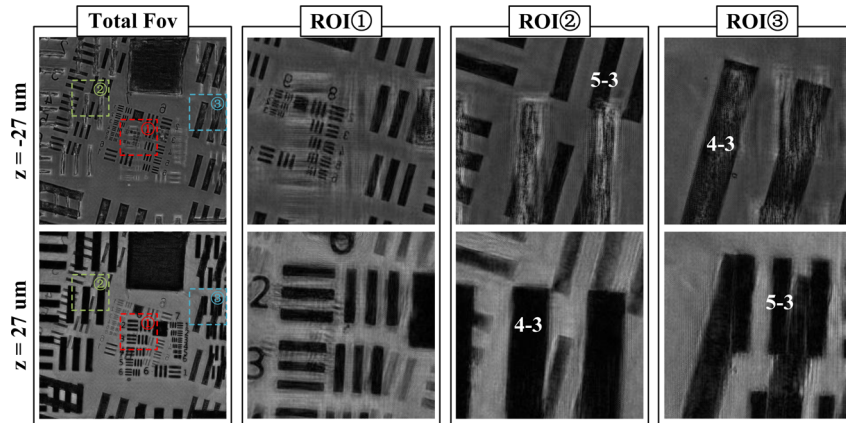


Fig. 6. Absorption reconstruction of a two-slice sample consisting of two resolution targets placed at different depths (-27 and $27 \mu\text{m}$), with one rotated laterally with respect to the other.

5. Experimental of biological sample

5.1. System setup

The raw data of *C.elegans* were obtained from Laura Waller's Computational Imaging Lab at UC Berkeley [44,64]. Since the diameter of *C.elegans* usually does not exceed $30 \mu\text{m}$, high-NA microscope objectives are used for high axial resolution. Consequently, all acquired images are brightfield images, and the raw data were applied HDR combined to calibrate the intensity between different illumination. Therefore, we do not use LICC for adaptive intensity calibration. Simultaneously, we divide the imaging depth of $24 \mu\text{m}$ or $30 \mu\text{m}$ into 121 layers, where the impact of backscattering becomes more pronounced. So we employ LIAC to estimate the attenuation of the forward propagation field caused by backscattering. The computing platform is aligned with the simulation. We employed PyTorch's mixed-precision training specifically for the *C.elegans* data. The total imaging time to complete 50 iterations of a $1200 \times 1200 \times 121$ voxels volume was 1.8 hours.

5.2. Experiment results

The captured raw data includes 120 images with illumination angles scanned on a spiral trajectory. Our biological experiment generates a high-resolution RI tomography of an adult *C.elegans* worm's head region ($1200 \times 1200 \times 60$ voxels), with voxel size $0.12 \times 0.12 \times 0.5 \mu\text{m}^3$. We note that 8-20 iterations are sufficient for visualizing the *C.elegans*'s 3D tomographic structure, taking 18-45 minutes (Visualization 1). MSNN can balance the time required with the quality of 3D imaging. Additionally, as we can utilize various propagation models by substituting the connections within the layers in MSNN, we also conducted experiments for applying the multi-layer Born model [53] in MSNN framework.

To quantitatively visualize the worm's 3D biological RI, we present RGB-colored cross-sectional images of all ROIs at various axial and lateral positions. In Fig. 7(a1)-(a3), we show two lateral slices through reconstruction volume at the axial position of $z = -7, -0.5, +5 \mu\text{m}$. The digestive system including the mouth, pharynx, and pharyngeal lumen, are identified. Behind the pharynx, we can observe the nematode's intestinal lumen, which contains many microsphere-like details. Fig. 7(a6) and (a7) show the axial cross-section of the pharyngeal lumen and pharynx, respectively. In Fig. 7(b1)-(b4), the lateral slice clearly show the pharyngeal lumen, pharynx, and other micron-sized structure at different axial position. We can also see the phenomenon

of *C.elegans* feeding in Fig. 7(c1)-(c4). The microspheres in front of the mouth of *C.elegans* are *E.coli* bacteria, a food source for the worm(3D RI Tomography results in [Visualization 2](#)).

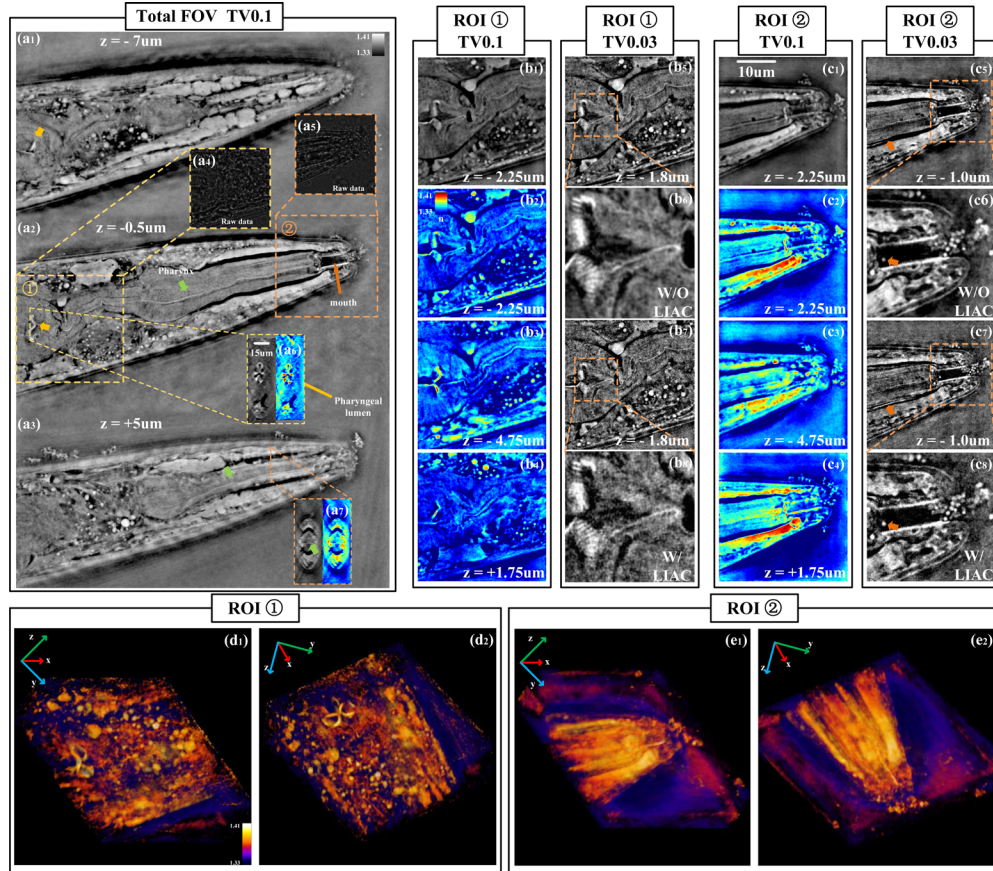


Fig. 7. RI tomography of a *C.elegans* worm's head. (a1)-(a3) Lateral slices through the 3D imaging volume at various axial positions. Major components of the digestive system are labeled. (a4)-(a5) The captured raw data illuminate by the positive incidence. Axial slice through the pharynx (a6), and axial slice through the pharyngeal lumen (a7) are shown. (b1)-(b4), (c1)-(c4) Lateral slices of the ROIs ① and ② in grayscale or RGB color at different axial positions. (b5)-(b8), (c5)-(c8) A comparison of experimental results with and without the implementation of LIAC. (d1)-(d2), (e1)-(e2) 3D reconstructions of the ROIs ① and ② from various viewing angles.

Due to the multi scattering effect inside *C.elegans*, backscattering at such long axial distances is still not negligible, so we conducted comparative experiments with and without LIAC. The results are shown in Fig. 7(b5)-(b8) and (c5)-(c8). As observed, the results with LIAC demonstrate enhanced clarity and higher contrast. During optimization, the implementation of LIAC also resulted in a smaller L1 loss compared to raw data.

Figure 7(d1)-(d2) and Fig. 7(e1)-(e2) show 3D visualizations from various viewing angles. To emphasize the 3D structure, we have obscured parts with an RI of less than 1.33. Usually, visualizing this morphology requires cross-sectional slicing and viewing by electron microscopy. We demonstrate that we can achieve non-invasive imaging to obtain 3D microscopic structures([Visualization 3](#)).

5.3. Validating the effectiveness of learnable intensity attenuation coefficients

To make the beam propagation model more closely resemble the actual propagation of the light field, we employ LIAC to predict the intensity of the backscattering field and attenuate the forward scattering fields, thereby enhancing the RI image quality. To analyze the impact of LIAC on the RI image quality for samples, we used MSNN to optimize 200 iterations without applying TV regularization. The total imaging time to complete 200 iterations of a $420 \times 420 \times 60$ voxels volume (The mouth of C.elegan) was 30 minutes. And the imaging results are shown in Fig. 8.

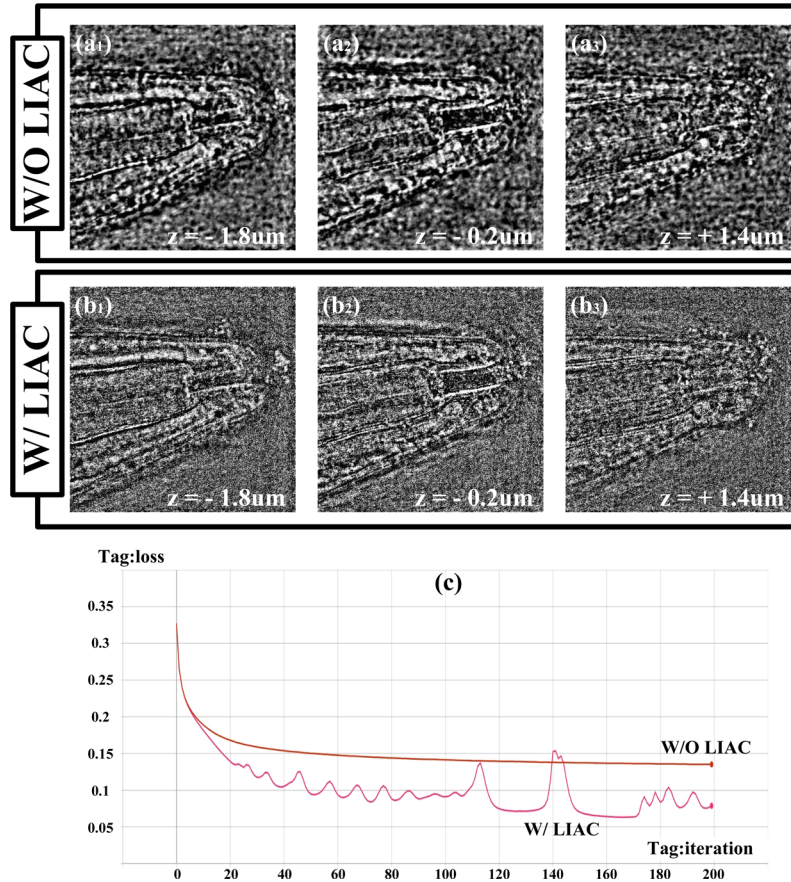


Fig. 8. Imaging results without TV regularization. (a1)-(a3) Lateral slices through the 3D reconstruction volume without the implementation of LIAC at various axial positions. (b1)-(b3) Lateral slices through the 3D reconstruction volume with the implementation of LIAC at various axial positions. (c) Loss curve of the optimizations

As observed, the loss curves indicate that the imaging results incorporating LIAC for predicted intensity align more closely with the captured raw data in the majority of iterations. Furthermore, the LIAC-enhanced results exhibit high-frequency noise which can be eliminated by many effective method, and we think that the high-frequency noise is caused by higher-order scattering fields. Imaging results without LIAC show more speckle-like noise, which would be difficult to remove, and we have a hypothesis that this noise is caused by not considering low-order scattering such as backscattering [65,66].

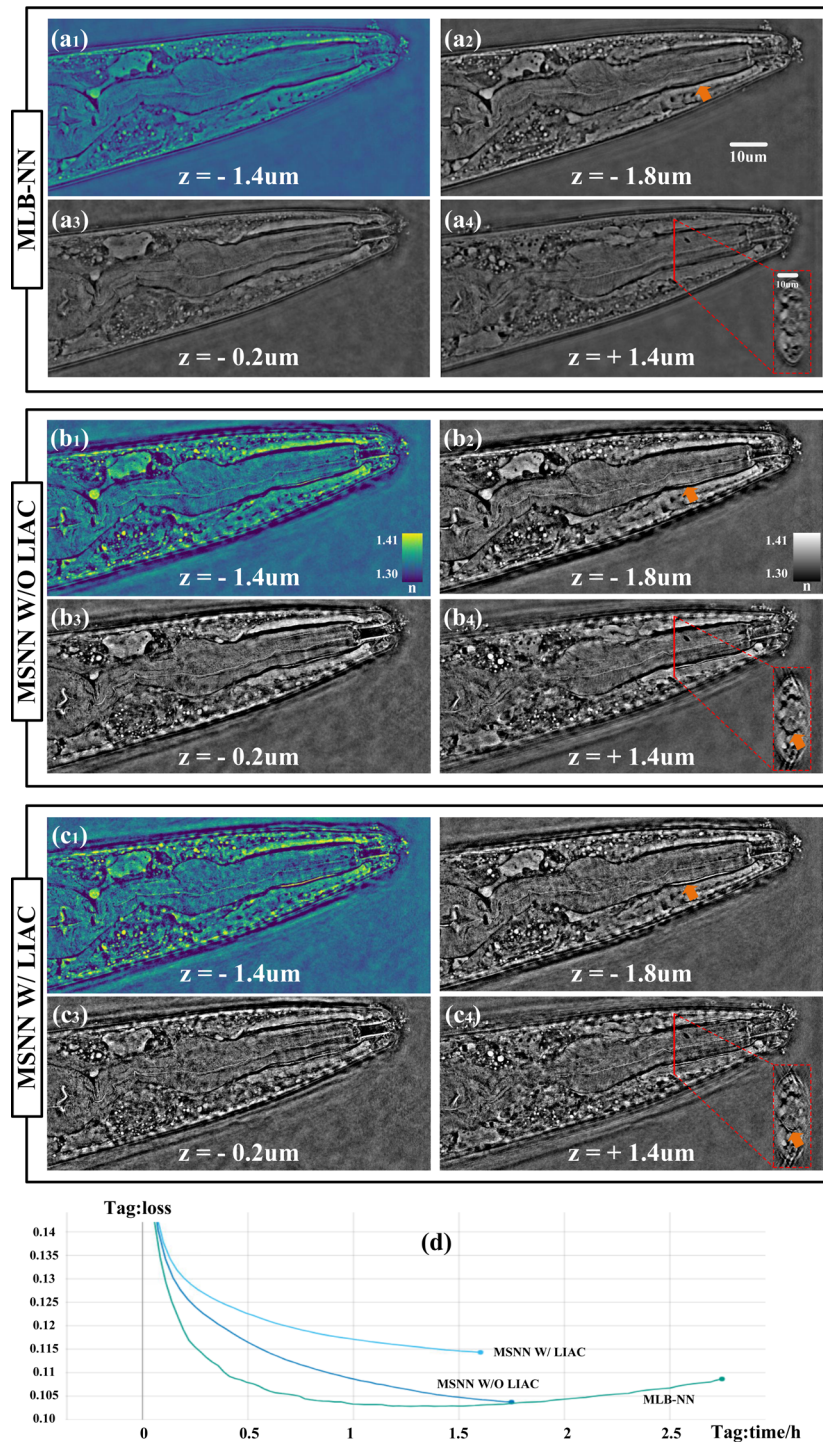


Fig. 9. The comparison of experimental results of a *C.elegan* worm's head with different methods. (a1)-(a4) Lateral slices through the sample imaging by neural network using multi-layer Born propagation model (MLB-NN). (b1)-(b4) Lateral slices through the sample imaging by the MSNN with LIAC. (c1-c4) Lateral slices through the sample imaging by the MSNN without LIAC. (d) Loss curves of the optimizations

5.4. Comparison of total field-of-view imaging with total variation regularization

Considering the challenge of separating information and noise in the reconstructed image without regularization, we present the total field-of-view reconstructed image using TV regularization with or without LIAC. Furthermore, the MSNN allows for the easy exchange of field-propagation models by defining the connections between layers. We also employ the multi-layer Born model within the Deep Neural Network (MLB-NN) for imaging the sample [53], as described in the methodology chapter of the text.

As shown in Fig. 9, the imaging utilizing MLB-NN shows a similar RI distribution to MSNN, albeit with lower contrast, and the loss converges at a slower pace. With the same level of TV regularization applied, the reconstructed image using MSNN with LIAC exhibits a quicker convergence speed and superior image granularity in the lateral slices and axial slices, which may reveal more details (labeled in Fig. 9) in the tomography.

6. Discussion and conclusion

An untrained physical based network termed as MSNN is proposed and demonstrated in this paper. It is an implementation of the paradigm of remodeling neural networks with optical principles. The scalability of MSNN provides a way to solve the constraints for RI tomography.

To enhance the clarity of RI tomography, darkfield images information is also very important. The implement of LICC can effectively calibrate the intensity between brightfield and darkfield, further, LICC can calibrate the intensity between different positions when using various LEDs (light sources) as illumination, the application of LICC has the ability to reduce hardware requirements and the need for preprocessing of raw data in RI tomography.

Furthermore, due to the complexity of multiple scattering, RI tomography is often limited to imaging samples with weak backscatter under the 1st Born or Rytov approximation. In the future, replacing the backscatter intensity estimation LIAC with a more accurate optical model may enable MSNN to account for higher-order scattering processes, showing promise in reconstructing samples with a broader RI range or extended axial distances.

In this paper, we have introduced MSNN, a neural network for efficient and accurate RI tomography. Compared to the widely used conventional deep learning methods, MSNN reconfigures the conventional CNN with the physical beam propagation model to confirm the authenticity of the 3D RI information. Moreover, it accelerates RI tomography compared to CNNs that employ physical method verification. MSNN introduces a novel paradigm for RI tomography, enabling to image biological samples (*C.elegans*) that are both thicker and exhibit higher scattering properties than were achievable with earlier methods. The implementation of LICC and LIAC also improved the robustness and imaging clarity and present its scalability.

Funding. National Natural Science Foundation of China (62275020).

Disclosures. The authors declare no conflicts of interest.

Data availability. We provide link to datasets and an open-source implementation of MSNN at GitHub repository available at [Code 1](#) [67].

References

1. D. T. Ross, U. Scherf, M. B. Eisen, *et al.*, "Systematic variation in gene expression patterns in human cancer cell lines," *Nat. Genet.* **24**(3), 227–235 (2000).
2. A. Rustom, R. Saffrich, I. Markovic, P. Walther, and H.-H. Gerdes, "Nanotubular highways for intercellular organelle transport," *Science* **303**(5660), 1007–1010 (2004).
3. M. Okuda, K. Li, M. R. Beard, L. A. Showalter, F. Scholle, S. M. Lemon, and S. A. Weinman, "Mitochondrial injury, oxidative stress, and antioxidant gene expression are induced by hepatitis c virus core protein," *Gastroenterology* **122**(2), 366–375 (2002).
4. R. Rizzuto, M. Brini, P. Pizzo, M. Murgia, and T. Pozzan, "Chimeric green fluorescent protein as a tool for visualizing subcellular organelles in living cells," *Curr. Biol.* **5**(6), 635–642 (1995).

5. E. Betzig, G. H. Patterson, R. Sougrat, O. W. Lindwasser, S. Olenych, J. S. Bonifacino, M. W. Davidson, J. Lippincott-Schwartz, and H. F. Hess, "Imaging intracellular fluorescent proteins at nanometer resolution," *Science* **313**(5793), 1642–1645 (2006).
6. S. W. Hell and J. Wichmann, "Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy," *Opt. Lett.* **19**(11), 780–782 (1994).
7. M. J. Rust, M. Bates, and X. Zhuang, "Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (storm)," *Nat. Methods* **3**(10), 793–796 (2006).
8. M. Bates, B. Huang, G. T. Dempsey, and X. Zhuang, "Multicolor super-resolution imaging with photo-switchable fluorescent probes," *Science* **317**(5845), 1749–1753 (2007).
9. S. Manley, J. M. Gillette, G. H. Patterson, H. Shroff, H. F. Hess, E. Betzig, and J. Lippincott-Schwartz, "High-density mapping of single-molecule trajectories with photoactivated localization microscopy," *Nat. Methods* **5**(2), 155–157 (2008).
10. S. W. Paddock, "Confocal laser scanning microscopy," *BioTechniques* **27**(5), 992–1004 (1999).
11. W. R. Zipfel, R. M. Williams, and W. W. Webb, "Nonlinear magic: multiphoton microscopy in the biosciences," *Nat. Biotechnol.* **21**(11), 1369–1377 (2003).
12. P. J. Keller, A. D. Schmidt, J. Wittbrodt, and E. H. Stelzer, "Reconstruction of zebrafish early embryonic development by scanned light sheet microscopy," *Science* **322**(5904), 1065–1069 (2008).
13. J. Sharpe, U. Ahlgren, P. Perry, B. Hill, A. Ross, J. Hecksher-Sørensen, R. Baldock, and D. Davidson, "Optical projection tomography as a tool for 3d microscopy and gene expression studies," *Science* **296**(5567), 541–545 (2002).
14. T. Alanentalo, A. Asayesh, H. Morrison, C. E. Lorén, D. Holmberg, J. Sharpe, and U. Ahlgren, "Tomographic molecular imaging and 3d quantification within adult mouse organs," *Nat. Methods* **4**(1), 31–33 (2007).
15. T. Correia, N. Lockwood, S. Kumar, *et al.*, "Accelerated optical projection tomography applied to in vivo imaging of zebrafish," *PLoS One* **10**(8), e0136213 (2015).
16. L. Tian and L. Waller, "3d intensity and phase imaging from light field measurements in an led array microscope," *Optica* **2**(2), 104–111 (2015).
17. R. Horstmeyer, J. Chung, X. Ou, G. Zheng, and C. Yang, "Diffraction tomography with Fourier ptychography," *Optica* **3**(8), 827–835 (2016).
18. R. Ling, W. Tahir, H.-Y. Lin, H. Lee, and L. Tian, "High-throughput intensity diffraction tomography with a computational microscope," *Biomed. Opt. Express* **9**(5), 2130–2141 (2018).
19. J. Li, A. Matlock, Y. Li, Q. Chen, C. Zuo, and L. Tian, "High-speed in vitro intensity diffraction tomography," *Adv. Photonics* **1**(06), 1 (2019).
20. C. Zuo, J. Sun, J. Li, A. Asundi, and Q. Chen, "Wide-field high-resolution 3d microscopy with Fourier ptychographic diffraction tomography," *Opt. Lasers Eng.* **128**, 106003 (2020).
21. S. Zhou, J. Li, J. Sun, N. Zhou, Q. Chen, and C. Zuo, "Accelerated Fourier ptychographic diffraction tomography with sparse annular led illuminations," *J. Biophotonics* **15**(3), e202100272 (2022).
22. J. Li, N. Zhou, J. Sun, S. Zhou, Z. Bai, L. Lu, Q. Chen, and C. Zuo, "Transport of intensity diffraction tomography with non-interferometric synthetic aperture for three-dimensional label-free microscopy," *Light: Sci. Appl.* **11**(1), 154 (2022).
23. S. Zhou, J. Li, J. Sun, N. Zhou, H. Ullah, Z. Bai, Q. Chen, and C. Zuo, "Transport-of-intensity Fourier ptychographic diffraction tomography: defying the matched illumination condition," *Optica* **9**(12), 1362–1373 (2022).
24. Y. Zewen, Z. Lu, L. Ning, W. Huijun, Z. Zhenxi, and Y. Li, "Progress of three-dimensional label-free quantitative imaging of refractive index in biological samples," *Chin. J. Laser* **49**, 0507201 (2022).
25. E. Kang, J. Min, and J. C. Ye, "A deep convolutional neural network using directional wavelets for low-dose X-ray ct reconstruction," *Med. Phys.* **44**(10), e360–e375 (2017).
26. K. H. Jin, M. T. McCann, E. Froustey, and M. Unser, "Deep convolutional neural network for inverse problems in imaging," *IEEE Trans. on Image Process.* **26**(9), 4509–4522 (2017).
27. N. Thanh, Y. Xue, Y. Li, L. Tian, and G. Nehmetallah, "Deep learning approach to Fourier ptychographic microscopy," *Opt. Express* **26**, 26470 (2018).
28. H. K. Aggarwal, M. P. Mani, and M. Jacob, "Modl: Model-based deep learning architecture for inverse problems," *IEEE Trans. Med. Imaging* **38**(2), 394–405 (2019).
29. Y. Sun, Z. Xia, and U. S. Kamilov, "Efficient and accurate inversion of multiple scattering with deep learning," *Opt. Express* **26**(11), 14678–14688 (2018).
30. Y. Li, Y. Xue, and L. Tian, "Deep speckle correlation: a deep learning approach toward scalable imaging through scattering media," *Optica* **5**(10), 1181–1190 (2018).
31. Z. Li, J. Sun, Y. Fan, Y. Jin, Q. Shen, M. Trusiak, M. Cywińska, P. Gao, Q. Chen, and C. Zuo, "Deep learning assisted variational Hilbert quantitative phase imaging," *Opto-Electronic Sci.* **2**(4), 220023 (2023).
32. Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature* **521**(7553), 436–444 (2015).
33. F. Xing, Y. Xie, H. Su, F. Liu, and L. Yang, "Deep learning in microscopy image analysis: a survey," *IEEE Trans. Neural Netw. Learning Syst.* **29**(10), 4550–4568 (2018).
34. E. Moen, D. Bannon, T. Kudo, W. Graf, M. Covert, and D. Van Valen, "Deep learning for cellular image analysis," *Nat. Methods* **16**(12), 1233–1246 (2019).
35. Y. Li, W. Jin, S. Qiu, D. Zuo, and J. Liu, "Color-coated steel sheet roof building extraction from external environment of high-speed rail based on high-resolution remote sensing images," *Remote Sens.* **15**(16), 3933 (2023).

36. A. Matlock and L. Tian, "Physical model simulator-trained neural network for computational 3d phase imaging of multiple-scattering samples," *arXiv*, arXiv:2103.15795 (2021).
37. G. Barbastathis, A. Ozcan, and G. Situ, "On the use of deep learning for computational imaging," *Optica* **6**(8), 921–943 (2019).
38. Z. Chao, F. Shijie, Z. Xiangyu, H. Jing, and Q. Chen, "Deep learning based computational imaging status challenges and future," *Acta Opt. Sin.* **40**(1), 0111003 (2020).
39. K. C. Zhou and R. Horstmeyer, "Diffraction tomography with a deep image prior," *Opt. Express* **28**(9), 12872–12896 (2020).
40. F. Wang, Y. Bian, H. Wang, M. Lyu, G. Pedrini, W. Osten, G. Barbastathis, and G. Situ, "Phase imaging with an untrained neural network," *Light: Sci. Appl.* **9**(1), 77 (2020).
41. R. Liu, Y. Sun, J. Zhu, L. Tian, and U. S. Kamilov, "Recovery of continuous 3d refractive index maps from discrete intensity-only measurements using neural fields," *Nat. Mach. Intell.* **4**(9), 781–791 (2022).
42. X. Lin, Y. Rivenson, N. T. Yardimci, M. Veli, Y. Luo, M. Jarrahi, and A. Ozcan, "All-optical machine learning using diffractive deep neural networks," *Science* **361**(6406), 1004–1008 (2018).
43. A. M. Maiden, M. J. Humphry, and J. M. Rodenburg, "Ptychographic transmission microscopy in three dimensions using a multi-slice approach," *J. Opt. Soc. Am. A* **29**(8), 1606–1614 (2012).
44. S. Chowdhury, M. Chen, R. Eckert, D. Ren, F. Wu, N. Repina, and L. Waller, "High-resolution 3d refractive index microscopy of multiple-scattering samples from intensity images," *Optica* **6**(9), 1211–1219 (2019).
45. J. Hu, S. Li, H. Xie, and Y. Shen, "Multi-slice ptychographic imaging with multistage coarse-to-fine reconstruction," *Opt. Express* **30**(12), 21211–21229 (2022).
46. U. S. Kamilov, I. N. Papadopoulos, M. H. Shoreh, A. Goy, C. Vonesch, M. Unser, and D. Psaltis, "Optical tomographic image reconstruction based on beam propagation and sparse regularization," *IEEE Trans. Comput. Imaging* **2**(1), 59–70 (2016).
47. U. S. Kamilov, I. N. Papadopoulos, M. H. Shoreh, A. Goy, C. Vonesch, M. Unser, and D. Psaltis, "Learning approach to optical tomography," *Optica* **2**(6), 517–522 (2015).
48. S. Jiang, K. Guo, J. Liao, and G. Zheng, "Solving Fourier ptychographic imaging problems via neural network modeling and tensorflow," *Biomed. Opt. Express* **9**(7), 3306–3319 (2018).
49. D. Yang, S. Zhang, C. Zheng, G. Zhou, L. Cao, Y. Hu, and Q. Hao, "Fourier ptychography multi-parameter neural network with composite physical priori optimization," *Biomed. Opt. Express* **13**(5), 2739–2753 (2022).
50. T. Chang, S. Shin, M. Lee, and Y. Park, "Computational approach to dark-field optical diffraction tomography," *APL Photonics* **5**(4), 040804 (2020).
51. K. Yanny, K. Monakhova, R. W. Shuai, and L. Waller, "Deep learning for fast spatially varying deconvolution," *Optica* **9**(1), 96–99 (2022).
52. S. Ruder, "An overview of gradient descent optimization algorithms," *arXiv* arXiv:1609.04747 (2016).
53. M. Chen, D. Ren, H.-Y. Liu, S. Chowdhury, and L. Waller, "Multi-layer born multiple-scattering model for 3d phase microscopy," *Optica* **7**(5), 394–403 (2020).
54. D. P. Kingma and J. Ba, "Adam: a method for stochastic optimization," *arXiv* arXiv:1412.6980 (2014).
55. C. Zheng, S. Zhang, G. Zhou, Y. Hu, and Q. Hao, "Robust Fourier ptychographic microscopy via a physics-based defocusing strategy for calibrating angle-varied LED illumination," *Biomed. Opt. Express* **13**(3), 1581–1594 (2022).
56. C. Zheng, S. Zhang, D. Yang, G. Zhou, Y. Hu, and Q. Hao, "Robust full-pose-parameter estimation for the led array in Fourier ptychographic microscopy," *Biomed. Opt. Express* **13**(8), 4468–4482 (2022).
57. Y. Shu, J. Sun, J. Lyu, Y. Fan, N. Zhou, R. Ye, G. Zheng, Q. Chen, and C. Zuo, "Adaptive optical quantitative phase imaging based on annular illumination Fourier ptychographic microscopy," *Photonix* **3**(1), 24 (2022).
58. P. Ewald, "Introduction to the dynamical theory of X-ray diffraction," *Acta Crystallogr., Sect. A: Cryst. Phys., Diff., Theor. Gen. Crystallogr.* **25**(1), 103–108 (1969).
59. J. M. Cowley, *Diffraction Physics* (Elsevier, 1995).
60. G. Zheng, C. Shen, S. Jiang, P. Song, and C. Yang, "Concept, implementations and applications of Fourier ptychography," *Nat. Rev. Phys.* **3**(3), 207–223 (2021).
61. S. Zhang, G. Zhou, C. Zheng, T. Li, Y. Hu, and Q. Hao, "Fast digital refocusing and depth of field extended Fourier ptychography microscopy," *Biomed. Opt. Express* **12**(9), 5544–5558 (2021).
62. J. W. Goodman, *Introduction to Fourier Optics* (Roberts and Company publishers, 2005).
63. G. Popescu, *Quantitative Phase Imaging of Cells and Tissues* (McGraw-Hill Education, 2011).
64. S. Chowdhury, "3d intensity-based ODT data," google, 2023, <https://drive.google.com/drive/folders/1RK6wbsvxxRJ1glrxGT7XR5EAIjnhCB0c>.
65. M. Azimi and A. Kak, "Distortion in diffraction tomography caused by multiple scattering," *IEEE Trans. Med. Imaging* **2**(4), 176–195 (1983).
66. K. Belkebir, P. C. Chaumet, and A. Sentenac, "Influence of multiple scattering on three-dimensional imaging with optical diffraction tomography," *J. Opt. Soc. Am. A* **23**(3), 586–595 (2006).
67. D. Yang, "Physics-based-3d-tomography-multi-slice-neural-network," Github 2023, <https://github.com/yang980130/Physics-based-3D-tomography-Multi-slice-neural-network>.